

VIRTUAL PAN/TILT CAMERA SYSTEM AND METHOD  
FOR VEHICLES

Statement Regarding Federally Sponsored Research or

5 Development:

This invention was developed in the course of work under U.S. government contract MDA972-02-9-0005. The U.S. government may possess certain rights in the invention.

10 Background of the Invention

This invention is directed to a camera system for vehicles for use with autonomous navigation systems (ANSs), and more particularly to a virtual pan/tilt camera system and method for use with ANSs for ground vehicles, more  
15 specifically manned ground combat vehicles.

Unmanned ground vehicles (UGVs) are valuable to the military involved in ground combat. The value of a UGV is increased when it is able to navigate autonomously, as it reduces manpower that must be deployed for a given military  
20 operation. Modern ANSs for ground vehicles use machine vision to detect obstacles and navigate paths around the obstacles. Current ANSs use pan/tilt stereo cameras on the outside of the

UGV so that the ANS computer can get a good view of the obstacles at close range.

ANSs used in modern unmanned ground vehicles operate using as sources of information: (1) a digital map of the terrain where the vehicle is operating that is precise to approximately a meter or better, and (2) redefined information that it deduces from the pan/tilt cameras used in the system. It also uses GPS coordinates to establish the vehicle's location. However, before the ANS directs the vehicle to begin moving, it may require a closer view of the area immediately surrounding the vehicle.

The autonomous navigation of ground vehicles is particularly difficult. To autonomously navigate over ground, detailed knowledge is required of objects down to a size that the vehicle can traverse. Therefore, ground navigation requires continuous detailed inspection of the area immediately ahead of the vehicle. For example, there may be obstacles such as trees, brush, rubble, boulders, and other objects that do not appear on the digital map. From a distance, the ANS can identify that a particular object is present, but may not be able to tell whether it is surmountable, i.e., whether the vehicle can travel over the object. To further identify and characterize such objects, the ANS causes the vehicle to navigate closer to the objects, and as it does so initiates suitable commands to the pan/tilt cameras to aim the cameras at the various nearby objects to gain better imagery of the objects. Objects that are classified as "impassable" are added to a dynamic map. The ANS selects an appropriate route, if such a route is available, and initiates commands to direct the vehicle between or around the impassable objects, and through or over the passable ones.

Pan/tilt cameras work well for an ANS and do not pose any excessive demands for most UGVs. However, problems occur with the use of ANSs on manned ground vehicles (MGVs). To enable MGVs to travel while a crew is resting or disabled, and to  
5 enable the vehicles to operate unmanned when facing extreme danger, it is required that the MGV be able to operate by remote control and navigate autonomously, such that the MGV must include an ANS. Because the current state of ANS technology requires a pan/tilt camera, a standard approach is  
10 to incorporate a pan/tilt camera into each MGV. However, the use of a pan/tilt camera with an MGV presents problems not associated with the use of such a camera with a UGV.

One such problem is that the pan/tilt camera increases vehicle weight and power needs which is of particular concern  
15 for MGVs where weight and power are severely constrained. Another problem is that the pan/tilt mount for the camera is difficult to hide from radar and infrared sensors, thereby compromising the low-observable characteristics needed for military vehicles to survive on the modern battlefield. This  
20 invention mitigates these problems by eliminating the need for the pan/tilt camera without requiring the development of new ANS technology and without incurring the costs and schedule delays associated with such development.

## 25 Summary of the Invention

In accordance with this invention the problems associated with the use of a pan/tilt camera with present ANS technology in an MGV are eliminated by eliminating the pan/tilt cameras on the vehicle and replacing them with a plurality of video  
30 infrared cameras. These video/IR cameras are normally included in an MGV to give the vehicle crew a view of the horizon in all directions. Each of the cameras is mounted in

a fixed orientation on the vehicle and covers a selected angular range of lateral view, preferably so that every point on the horizon is viewed by two different cameras such that the cameras' fields of view overlap enabling every point on the horizon to be viewed even when a camera is disabled.

In prior art, MGVs are equipped with such cameras, and computers and software in the MGV continuously receive data from all of the cameras, store the data in video frame buffers, and stitch the imagery together into a single, seamless panoramic view. In some MGVs, each crew member of an MGV can select which subset of the view to display on a particular computer console. A display computer maps the selected view to a rectangular region of a video frame buffer, and data from that region is sent to the particular computer console each time a new video frame comes into the buffer. A crew member of the MGV can adjust the selected view to the right or left, up or down, or zoom in or out.

In accordance with the invention, computers and software map pan/tilt commands for individual cameras from the MGV's ANS to regions of the MGV's video frame buffers, and feed appropriately transformed data from the video frame buffers to the ANS. Video data from two different cameras with overlapping fields of view are used to simulate the view of stereo cameras.

Thus, the present invention eliminates the need for pan/tilt cameras and the disadvantages associated with their use on MGVs, by using a plurality of fixed cameras already included on modern MGVs, each camera covering a selected angular range of lateral view with the imagery from the cameras stitched together to produce a single, seamless panoramic view as may be selected by a crew member. With the present invention, the MGV does not suffer penalties to its

radar and infrared signatures as would occur with the use of pan/tilt cameras and their associated mounts. Moreover, the present invention may be used with present ANS technology thereby eliminating the substantial costs that would be associated with developing a new type of ANS. The present invention reduces weight and power requirements as compared to the use of pan/tilt cameras.

Further features and advantages of the present invention, as well as the structure and operation of various embodiments of the present invention, are described in detail below with reference to the accompanying drawings.

#### Brief Description of the Drawings

Figure 1 is a schematic block diagram of a preferred embodiment of the invention;

Figure 2 is a diagram illustrating image distortion;

Figure 3 is a diagram illustrating a geometrical transformation in accordance with a preferred embodiment of the invention; and

Figure 4 is an illustration of image transformation in accordance with a preferred embodiment of the invention.

#### Detailed Description of the Preferred Embodiments

With reference to the drawings, there is shown a preferred embodiment of a virtual pan/tilt camera system 10 in accordance with the present invention for use with a manned ground combat vehicle 12 as illustrated in Figure 3. The system generally comprises a fixed camera video data input 14, a reformatting network section 16, and a display system section 18. The fixed camera video data input 14 further comprises a plurality of video cameras 20, each of which is mounted in a fixed position on the vehicle 12. In this

embodiment eight such cameras are shown, but it is to be understood that a fewer or greater number of cameras may be used. A suitable camera is the Hitachi HV-C20 or the Sony DXC 9000. Each camera outputs video data 22 representing the  
5 images viewed by that camera to a frame buffer 24, there being eight such frame buffers, one for each video camera. Each frame buffer stores at addresses within the buffer the video data it receives from its respective camera.

An ANS 26 of the same type used for autonomous navigation  
10 of unmanned ground vehicles that use pan/tilt stereo camera pairs is programmed to operate as if it were in a system using a pan/tilt stereo camera pair. Thus, as previously explained, the ANS is a pre-existing system that performs stereo image processing, and includes appropriate hardware and software to  
15 autonomously navigate the vehicle as is known in the art. An example of such a system is an ANS incorporated in the experimental unmanned vehicle developed by General Dynamics Robotic Systems and the National Institute of Standards and Technology for the Department of Defense's Demo III program.  
20 What the ANS needs in order to perform its required functions is two video input streams from a pan/tilt stereo camera pair that it can control. With the system 10 of the present invention, there are no pan/tilt cameras. Instead, the system 10 functions to deliver two video input streams generated from  
25 fixed cameras, but which have been reformatted to appear to have come from a pan/tilt stereo pair.

Thus, the ANS determines the best angle at which to point what would be a pan/tilt stereo pair to get a view of an obstacle, and outputs appropriate pan/tilt commands 28 to the  
30 reformatting network system 16. The system 16 comprises a Translate to Az-El component 30, an Az-El to Address Translation component 32, and a Video Formatting component 34.

The pan/tilt commands 28 from the ANS 26 are input to the Translate to Az-El component 30 which converts the pan/tilt commands into signals representing the azimuth and elevation for each pan/tilt camera pair as if such cameras were present in the system. The azimuth and elevation output signals 36 from the Translate to Az-El component 30 are input to the Az-El to Address Translation component 32 which generates address signals 40 which are input to each of the video frame buffers 24. The Az-El to Address Translation component 32 maps the azimuth and elevation signals, as commanded by the ANS, to addresses in the video frame buffers 24, thereby determining which of the frame buffers 24 and which data stored in each buffer to access for generating each section of the image.

Video data 50 from the specified addresses, as determined by the Az-El to Address Translation component 32, in each of the frame buffers 24 are input to the Video Formatting component 34 which stitches together images from adjacent frame buffers, if needed, and reformats the images to correct the image geometry, made necessary because the fixed cameras 20 do not actually point in the direction specified by the pan/tilt commands 28 from the ANS 26. An image formed to the side of an image plane of a camera 20 is distorted compared to the same image formed at the center of the image plane as illustrated in Figure 2. Figure 2A illustrates an image, represented by arrows 60, 62, centered on a camera axis 64, only minimally distorted at the focal plane 66 as shown by the arrows 68, 70 of substantially equal length. Figure 2B shows the same image, represented by arrows 60, 62, centered off the camera axis 64, distorted at the focal plane 66 as shown by the arrows 72, 74 of unequal length in the image plane.

This is further illustrated and explained with reference to Figures 3A and 3B. Figure 3A illustrates a ground combat

vehicle 12 having two pan/tilt cameras A and B, each pointing directly, or nearly directly, to an image represented by arrows 80, 82. That is, the camera axes 84, 86 of the cameras A and B are aligned, or nearly aligned, with direct paths, as illustrated by lines 90, 92, between the cameras and the image. Figure 3B illustrates the same ground combat vehicle 12, but instead of two pan/tilt cameras as in Figure 3A, the vehicle is equipped with fixed video cameras 20. In this preferred embodiment there are eight such fixed cameras equally circumferentially spaced about a common axis 94 such that collectively they encompass a 360° panoramic view about the vehicle. Thus, the camera axes 96 of the cameras 20 point radially from the central axis 94 and are 45 degrees apart. Each of the cameras covers at least a 90° panoramic view centered about its axis such that adjacent cameras overlap the field of view between them. This way the entire 360° field of view is covered even when one camera is disabled. Thus, two such cameras a and b, viewing the same image as the cameras A and B of Figure 3A, have their camera axes 96 substantially out of alignment with the direct paths, as shown by lines 104, 106 between the cameras a and b and the image.

For the ANS to perform properly, it must receive two video streams at its inputs 110, 112 that are substantially the same as the video streams it would receive if the image was being viewed by two pan/tilt cameras as in Figure 3A. However, the video signals generated by the fixed cameras 20 and stored in the video frame buffers do not produce the same video stream as two pan/tilt cameras, because the fixed cameras do not actually point in the direction specified by the pan/tilt commands 28 from the ANS. Therefore, in accordance with the invention, the Video Formatting component 34 reformats the images to correct the image geometry.



The functions performed by the Video Formatting component 34 can be implemented as hardware or software. Most preferably, the Video Formatting component 34 is implemented using a programmable image processing card which may be a purchased off-the-shelf item. As illustrated in Figures 3A and 3B, the software on the image processing card transforms the image from a fixed camera's focal plane to the image that would be seen on a pan/tilt camera's focal plane, assuming that the pan/tilt camera was pointed in the direction that the ANS commanded. The camera's focal plane may be a charged coupled device or any device that transforms light images into electrical signals. An image in a video camera is defined by luminous intensity as a function of x and y locations on the CCD, i.e., brightness at each pixel. This may be described as  $i=f(x,y)$ . To correctly transform the image from a fixed camera to make it appear like an image from a pan/tilt camera, the software performs certain steps. In describing these steps, reference is made to Figures 4a-4g which illustrate an image, shown as a tree figure, on the focal plane of a camera (fixed or virtual).

1) convert pixel indices (rowfixed, columnfixed) for the fixed camera focal plane to physical locations on the fixed camera's focal plane, i.e., (xfixed, yfixed).

2) convert the image from  $i=f_1(xfixed, yfixed)$ , i.e. intensity as a function of x and y locations on the fixed camera's focal plane, to intensity as a function of Az-El coordinates relative to the fixed camera's optical axis, i.e.  $i=f_2(Azfixed, Elfixed)$ . This step yields a curved focal plane as shown in Figure 4c.

3) rotate the Az-El coordinates from step 2 to the appropriate Az-El coordinates relative to the commanded optical axis of the virtual pan/tilt camera, i.e.

Azvirtual=g1(Azfixed, Elfixed) and Elvirtual=g2(Azfixed, Elfixed).

4) convert the resulting Az-El image,  $i=f3(\text{Azvirtual}, \text{Elvirtual})$ , to intensity as a function of  $x$  and  $y$  on the virtual pan/tilt camera's focal plane, i.e.,  $I=f4(x\text{virtual}, y\text{virtual})$ . The result of this step is a four-sided polygon that is not an exact parallelogram as it has been transformed to a distorted form from the initial parallelogram of Figures 4a and 4b.

5) convert the physical locations  $(x\text{virtual}, y\text{virtual})$  to pixel indices, yielding  $I=f5(\text{rowvirtual}, \text{columnvirtual})$ .

6) Interpolate the pixel values. In general, the center of a pixel in the fixed camera will not map directly to the center of a pixel in the virtual pan/tilt camera, so the  $(\text{rowvirtual}, \text{columnvirtual})$  values computed in step 5 will be floating point numbers, not integers. Therefore, the Video Formatting component interpolates the intensity values given by  $I=f5(\text{rowvirtual}, \text{columnvirtual})$ , where the values of  $(\text{rowvirtual}, \text{columnvirtual})$  do not align with the centers of pixels, to estimate the intensity values at the center of each pixel.

Steps 5 and 6 yield a focal plane shape that is the same as the shape from step 4. Step 5 yields a set of rows and columns that are wider and thicker at the front of the view and the rows and columns line up with the edges of the shape. Step 6 yields a set of rows and columns that have uniform width, such that they do not line up with all edges of the shape.

The algorithms used to solve each of these steps are within the skill of those skilled in the arts of optics and image processing.

The stitched and reformatted video signals from the Video Formatting component 34 are input as two video streams 110, 112 to the ANS 26 which outputs the pan/tilt commands 28 as previously described, which, if pan/tilt cameras were present, would control the direction of the pan/tilt cameras. However, in accordance with the present invention, the pan/tilt commands rather than controlling pan/tilt cameras are input to the Translate to Az-El component 30 which converts the pan/tilt commands into the azimuth and elevation that each pan/tilt camera would view if it received those commands. As previously explained, the azimuth and elevation values from the Translate to Az-El component 30 are input to the Az-El to Address Translation component 32 which performs the dual functions of mapping the azimuth and elevation signals to addresses in each frame buffer 24, and determining which frame buffer of the frame buffers 24 to use for each section of the image. The algorithms for performing these functions are known in the art.

The azimuth and elevation values from the translate to Az-El component 30 also are input to the Video Formatting 34 which uses these values to determine the amount of distortion, i.e., the off-axis angle, and which camera to use.

The display system section 18 comprises an Image Stitcher 120, a Panoramic View Buffer 122, an Az-El to Address Translation component 124, an Az-El View Selector(s) 126, and image displays 128. The display system section 18 is known in the art for processing video data to display selected images on displays in manned vehicles for viewing by the crew members. The image stitcher 120 receives video data from the video frame buffers 124, performs various transformations on selected subsets of the video data, and sends the resultant subframe video data to the Panoramic View Buffer 122. The

buffer 122 contains image data representing a 360° wide, by approximately 135° high, external view from the vehicle 12. By the Az-El View Selector(s) 126, a crew member selects an azimuth and elevation for viewing, and the Az-El View  
5 Selector(s) 126 outputs azimuth/elevation data, in accordance with those selections, to the Az-El to Address Translation component 124, which converts the azimuth/elevation data to a range of memory addresses in the Panoramic View Buffer 122. The image data from the range of memory addresses are output  
10 from the Panoramic View Buffer 122 to the displays 128 via the Az-El to Address Translation component 124 to display the selected portion of the panoramic view.

Video from two different cameras with overlapping fields of view is used to simulate the view of stereo cameras.

15 Selecting different ranges of addresses of each of the video frame buffers 24 simulates the effect of panning left or right and tilting up or down as with pan/tilt cameras.

Some ANS systems include a status interface that allows the ANS to check the functionality and position of each  
20 pan/tilt camera and mount. For such systems, if the pan/tilt cameras were not present, the ANS would interpret the absence of pan/tilt cameras as a malfunction and would not operate properly. Thus, the present invention includes a simulator that simulates the status of the pan/tilt cameras and mounts.

25 The simulator receives status inquiries from the ANS and returns appropriate data. For example, if queried for the current Az-El of the pan/tilt camera, the simulator returns the last commanded Az-El (because that is what a real pan/tilt camera would do if operating properly). If queried for  
30 operational status, it returns an "on" if the fixed cameras are turned on, and an "off" if they are turned off.

Thus, the virtual pan/tilt camera system and method of the present invention, particularly adapted for use with MGVs, eliminates problems associated with the use of pan/tilt cameras, including the increased vehicle weight and power  
5 needs for such cameras and the difficulty with hiding such cameras from radar and infrared sensors thereby compromising the low-observable characteristics needed for military vehicles. The present invention eliminates the need for pan/tilt cameras and their associated problems, providing a  
10 system and method utilizing fixed cameras with known ANS technology.

The ANS used with the present invention performs stereo imaging processing and autonomously navigates the ground vehicle in the same way it performs these functions when used  
15 in a pan/tilt camera system. The ANS, when used with pan/tilt cameras, receives two video data streams from a pair of pan/tilt cameras for the ANS to use to control the cameras by performing stereo image analyses. In accordance with the present invention, the same ANS receives two video data  
20 streams that "appear to" come from two pan/tilt cameras, when in fact they are video data streams that have been reformatted from fixed cameras.

While the present invention has been described by reference to specific embodiments and specific uses, it should  
25 be understood that other configurations and arrangements could be constructed, and different uses could be made, without departing from the scope of the invention as set forth in the following claims.